AO system Design: Astronomy

Olivier Guyon (Subaru Telescope) guyon@naoj.org

This lecture:

Will not discuss detailed optical designs, mechanical designs, hardware choices, computer algorithms
(often too specific to some AO systems)

Main goal is to explore AO strategies, compare them, understand how/why/when they work or don't work

This course won't teach you how to build an AO system, but it will help you figure out what kind of AO system you might build for a specific application & what kind of problems will need to be solved
Need to think simultaneously about science requirements, WFS strategy and hardware choices.

Outline

1. Main challenges / error budget terms in astronomical AO systems

How to design an AO system which meets science requirements & reduces overall error budget ?

- The answer strongly depends on the science objective
  - Solar AO, Extreme AO, MOAO, GLAO, LGS vs. NGS ...

2. Wavefront sensing strategy

- AO guide star:
  - LGS, NGS ? Multiple sources ? Sensing wavelength ?
- Choosing the right Wavefront Sensor

3. From photons to DM commands: making it all work nicely together

1. Main challenges / error budget terms in astronomical AO systems

How to design an AO system which meets science requirements & reduces overall error budget ?
- The answer strongly depends on the science objective
  - Solar AO, Extreme AO, MOAO, GLAO, LGS vs. NGS ...

2. Wavefront sensing strategy

- **AO guide star:**
  - **LGS, NGS ? Multiple sources ? Sensing wavelength ?**
- **Choosing the right Wavefront Sensor**

3. From photons to DM commands: making it all work nicely together

Fundamental AO challenges :

1 Fitting error

2 Speed

3 Limited # of photons

4 AO guide “star” size & structure, sky background

5 Non-common path errors
- chromaticity
- cone effect (LGS) & anisoplanetism

6 Calibration, nasty “practical” things
- vibrations, instabilities between control loops
- DM hysteresis / poor calibration (generally not too serious in closed loop)

1. Fitting error

Wavefront errors from atmospheric turbulence

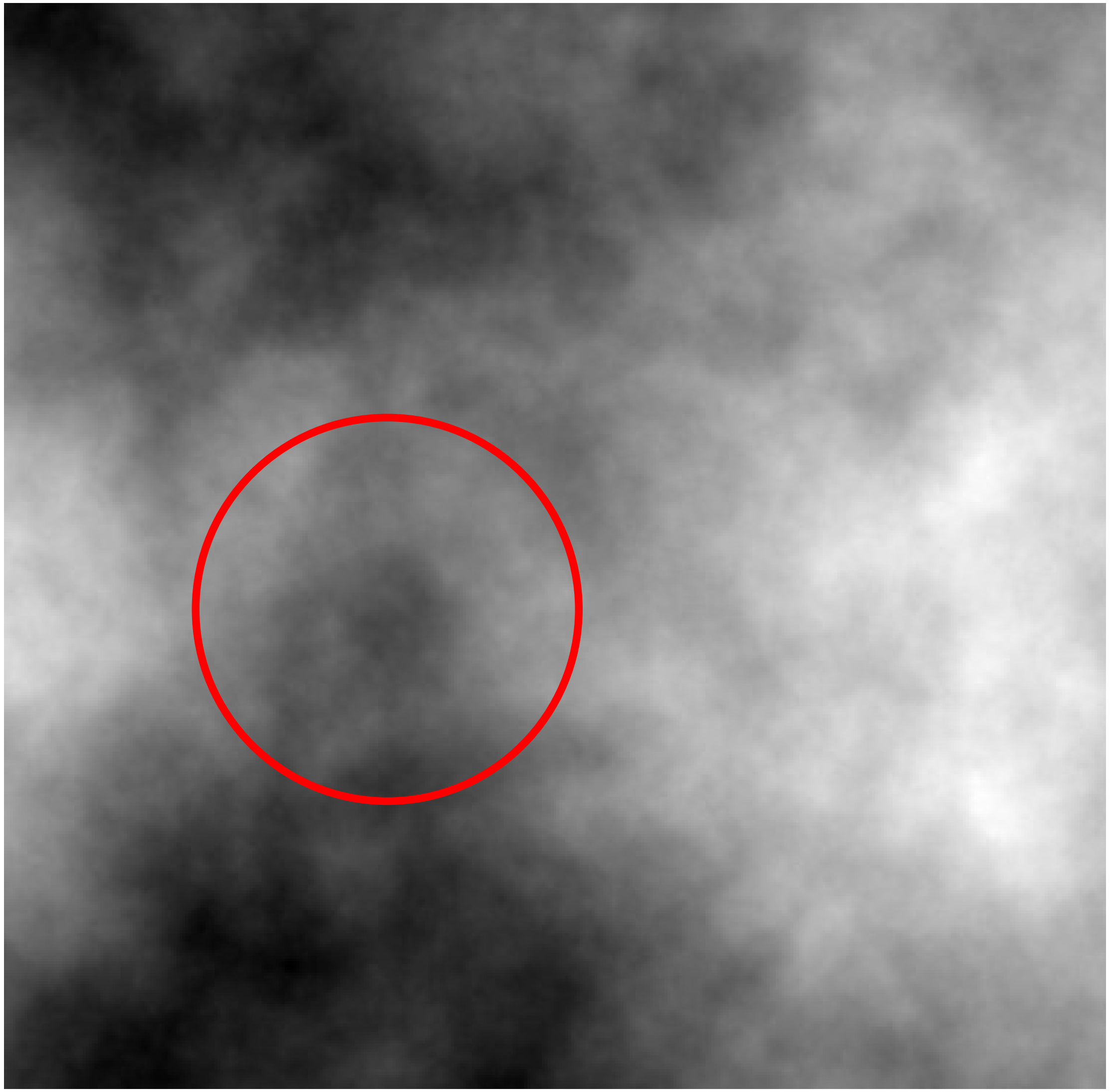
$$\sigma^2 = 1.03\,(D/r_0)^{5/3}$$

+ Vibrations, telescope guiding errors

+ Aberrations from optical elements
(primary mirror, large number of small mirrors)

+ DM shape at rest

Kolmogorov turbulence



1. Fitting error

Need enough stroke on the actuators

$\sigma^2 = 1.03 \, (D/r_0)^{5/3}$

Larger D -> more stroke needed

(also: faster system -> more stroke needed)

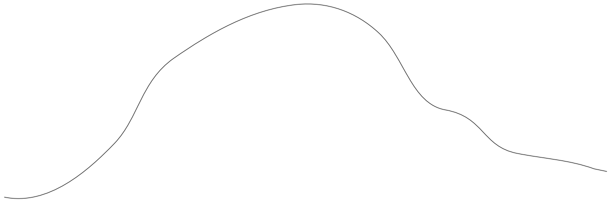
Most of the power is in tip-tilt:

It is helpful to have a dedicated tip-tilt mirror, or mount the DM on a tip-tilt mount

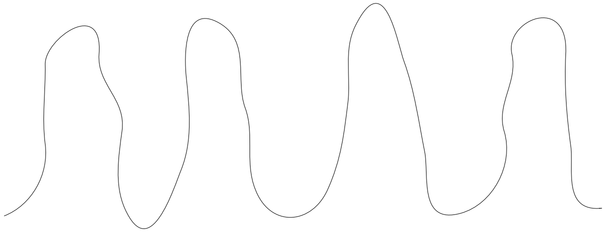
On many DMs, interactuator stroke < overall stroke

DM stroke needs to be looked at as a function of spatial frequency

eg: in a curvature DM, radius of curvature decreases as the number of actuators increases



Is easier than



1. Fitting error

Need enough actuators to fit the wavefront

D = telescope diameter

N = number of actuators

$d^2 = D^2/N$ = actuator size

If we assume each actuator does perfect piston correction (but no tip/tilt):

$\sigma^2 = 1.03\ (d/r_0)^{5/3} = 1.03\ (D/r_0)^{5/3}\ N^{-5/6}$

If we assume continuous facesheet (see Hardy, Roddier),

$\sigma^2 \sim 0.3\ (D/r_0)^{5/3}\ N^{-5/6}$

$D = 8$ m

$r_0 = 0.8$ m (0.2 m in visible = 0.8 m at 1.6 micron)

Diffraction limit requires $\sim N = 24$

In fact, exact DM geometry & influence functions are needed to estimate fitting error

2. Speed

assuming pure time delay t

$\sigma^2 = (t/t_0)^{5/3}$

t_0 = coherence time “Greenwood time delay” (see Hardy)

$= 0.314\ r_0/v$

v = 10 m/s

r_0 = 0.15 m (visible) 0.8 m (K band)

t_0 = 4.71 ms (visible) 25 ms (K band)

sampling frequency ~ 10x bandwidth

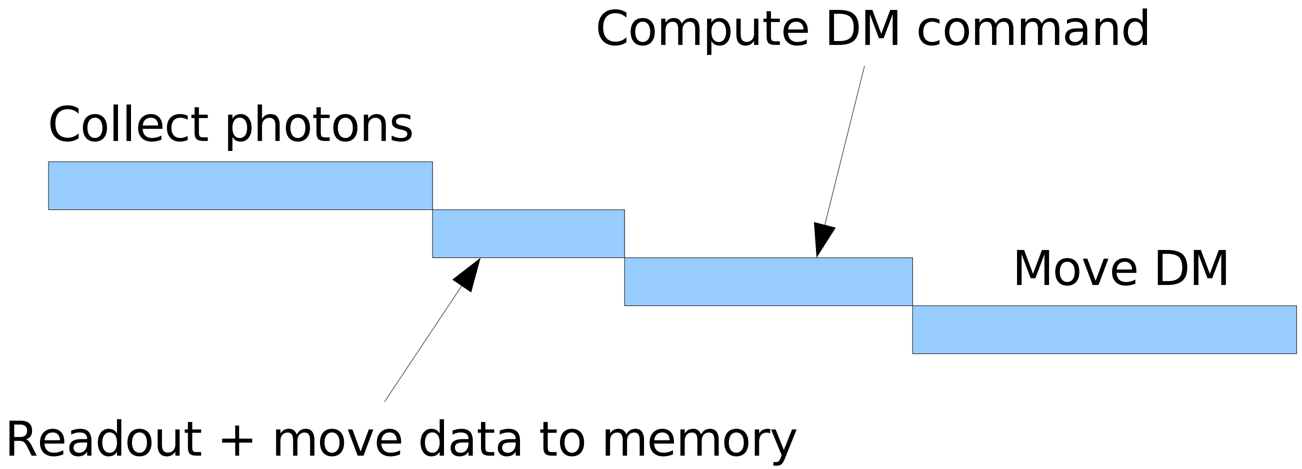
for “diffraction-limited” system (1 rad error in wavefront):
400 Hz for K band

for “extreme-AO” system (0.1 rad error):
6 kHz for K band

-> High speed means fewer photons / sample
need high SNR in WFS (optimal use of photons)

-> need fast hardware (see below)
- DM: good time response, low vibration
- Detector: fast readout / low readout noise
- computer, software & electronics

-> Clever, predictive control can help a lot
"anything that could be predicted should be !"



3. Limited # of photons from stars (per unit of time)

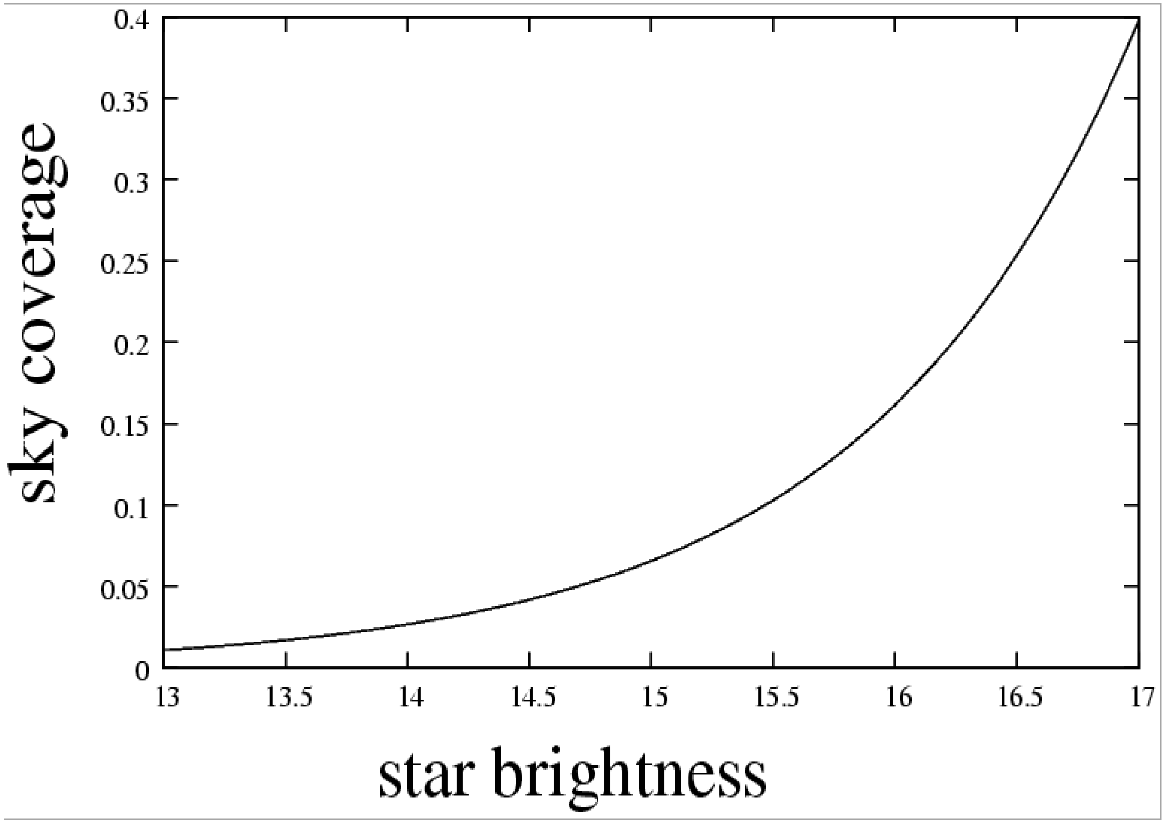
mV=15 -> 400 ph/ms on 8m pupil in 0.5 micron band
& 20% efficiency

Example 1: General purpose NGS system

Goal: achieve diffraction limited performance over much of the sky

Star brighter than mV density
~ 9e-4 exp(0.9 mV) per sq. deg (galactic pole)
ref: Parenti & Sasiela, 1994

Within a 20” radius:



mV=8 -> 2.5e5 ph/ms on 8m pupil in 0.5 micron band & 20% efficiency

Example 2: Extreme-AO system
Goal: Achieve exquisite wavefront correction on selected bright stars

Running speed = 5 kHz (see speed section before)
2000 actuators

25 photons / actuators / sampling time
6 photon / pixel if 2x2 Shack Hartmann cells are used

with no readout noise, ~ 0.2 rad phase error per actuator at best.

Limited # of photons will push system design into:

-> high efficiency WFS: good at converting OPD error into signal
(if possible, choose shorter wavelength)

-> high throughput (fewer optics), good detector (low readout noise)

-> WFS which works in broad band for NGS

-> bright laser for LGS, small angular size LGS

-> multiple guide stars

4 AO guide “star” size & structure, sky background

Extended targets means lower WFS efficiency and/or WFS failure

This problem is very WFS-dependent (some WFSs cannot deal with extended sources)

- Laser guide star is typically 1” or more, and elongated
- NGS: atmospheric refraction can be serious
    -> Atmospheric Dispersion Compensator (ADC) is essential
- frequent problem in Solar system observations
- double stars

Sky background:
for faint guide stars, moonlight is a concern

5 Non-common path errors

- anisoplanatism

  Due angular separation between guide star and science target, guide star WF is different from science wavefront

  -> minimize distance between guide star & science field

  -> use several guide stars & perform tomographic rec.

  -> if FOV is needed, use several guide stars (NGS or LGS)

- chromaticity

  AO correction is optimal for WFS wavelength, not for science wavelength

- cone effect (LGS)

  -> tomographic reconstruction

- instrumental non-common path errors

  Due to optics in WFS only or in science camera only

  -> may need to be measured (phase diversity works well for this) and offset to AO loop

6 Calibration, nasty "practical" things

- vibrations, instabilities between control loops
  - -> good mechanical design
  - -> beware of cryocoolers (pumps), fans, wind, telescope mounts, tip-tilt mirrors

- DM hysteresis / poor calibration (generally not too serious in closed loop)

Science wavelength choice:

IR is “easy”, visible is “very very hard”

Things that get worse as lambda gets small:

- r0 gets small: more actuators needed

  $r_0 \propto \lambda^{6/5}$ -> $N \propto \lambda^{-12/5}$

- speed gets high ($\tau_0 = 0.314\ r_0/v$) -> $\tau_0 \propto \lambda^{6/5}$

- anisoplanatism gets small (FOV, sky coverage go down)

  $\theta_0 \propto \lambda^{6/5}$

- chromaticity gets worse (refraction index of air varies more in visible than near-IR)

- instrumental non-common path errors get more serious

But diffraction limit is small in visible

Number of actuators should be very carefully chosen

Resist temptation of having more actuators than needed:

Systems with too many actuators are:

- not very sensitive (don't work well on faint stars)
- demanding on hardware, more complex & costly
- less tolerant (alignment, detector readout noise...)

There is usually little motivation to have much more than ~1 actuators per r0.

Exception:
Extreme-AO, where actuator # driven by field of view.

1. Main challenges / error budget terms in astronomical AO systems

How to design an AO system which meets science requirements & reduces overall error budget ?
- The answer strongly depends on the science objective
  - Solar AO, Extreme AO, MOAO, GLAO, LGS vs. NGS ...

2. Wavefront sensing strategy
- **AO guide star:**
  - **LGS, NGS ? Multiple sources ? Sensing wavelength ?**
- **Choosing the right Wavefront Sensor**

3. From photons to DM commands: making it all work nicely together

Choosing the wavefront sensing strategy is the most fundamental step in the design of an AO system

Which science ? AO system requirements ?
- Field of view ? Sky coverage ? Science wavelength ? Wavefront quality ?
  - Optimize for bright stars (ExAO), faint stars ?

-> 3 BIG questions need to be answered:
- Where to get the wavefront measurement ?
  - Natural guide star / Laser guide star ?
  - How many guide stars, where ?
- Which wavefront sensor(s) concept to adopt ?
  - It is important to get the fundamental physics of wavefront sensing right.
- How many elements in the WFS ?
  - This will drive the DM technology, computers etc...

Where to get the wavefront measurement ?

(1) Are there suitable natural guide stars ?

If not -> **Laser Guide Star (LGS)**

which laser ?

- Rayleigh
  low altitude (few km) Rayleigh scattering
  same process makes the sky blue
  works better at shorter wavelength
- Sodium
  excitation of sodium layer at 90 km
- Polychromatic Sodium (not quite ready yet)
  excitation of sodium layer to produce LGS
  in 2 wavelengths -> solves Tip/Tilt problem

LGS allows large (>50%) sky coverage

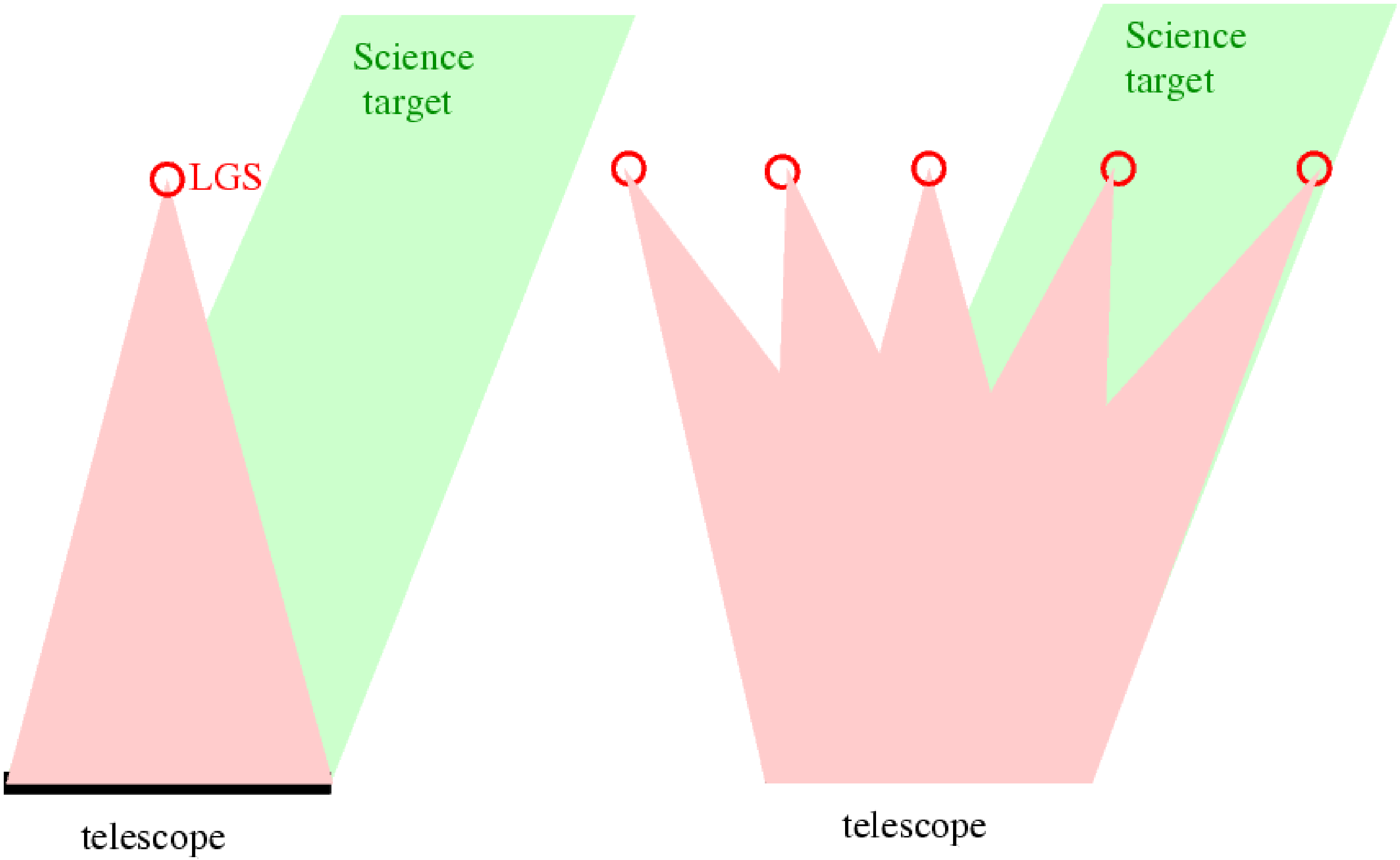
Where to get the wavefront measurement ?

(2) Need several guide stars ?

(for field of view, tomography ?)

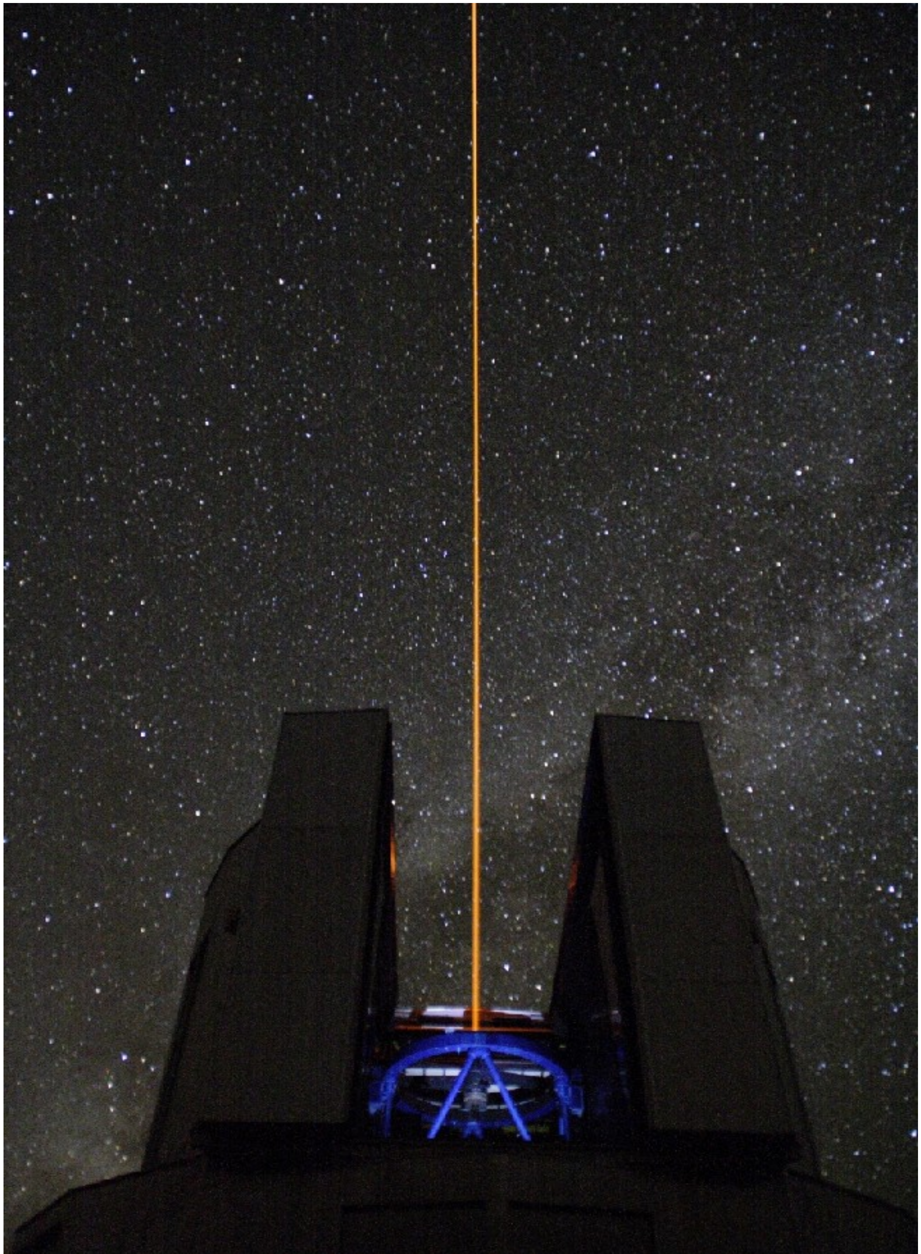
Multiple LGS ?

Multiple NGS ?



Some challenges of LGS AO

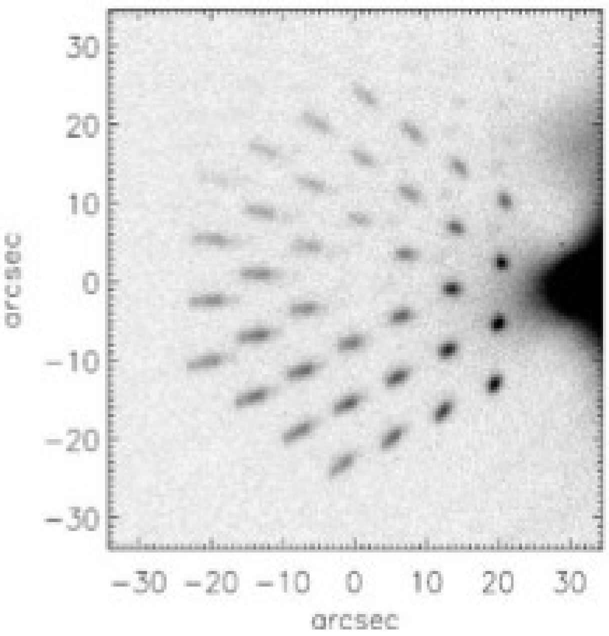
Cone effect due to finite altitude of LGS (90km sodium, few km for Rayleigh)
-> can be solved by using several lasers and tomography



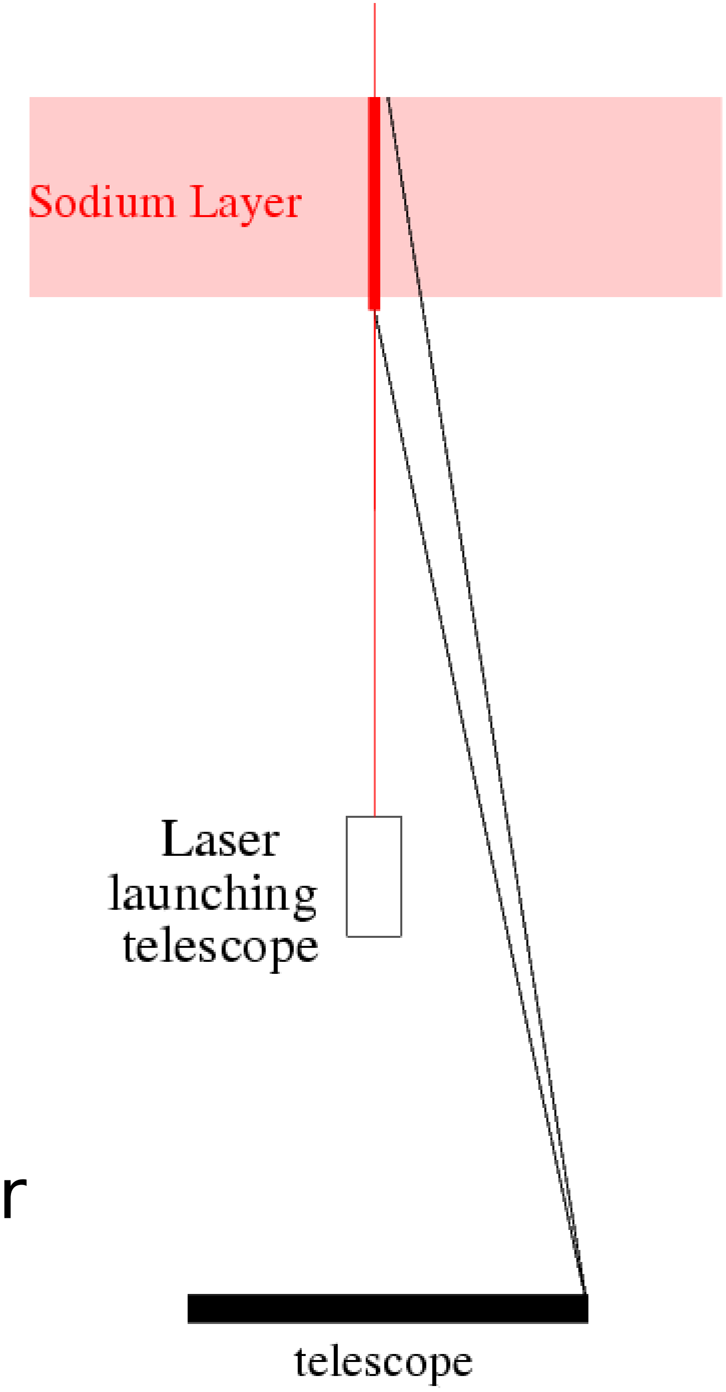
Tip/Tilt & Focus sensing
Upstream & downstream paths are the same: tip/tilt not seen
Sodium layer altitude not fixed: LGS focus info is incomplete (can be used to sense fast focus)
-> **Still need NGS(s) for tip/tilt & Focus**
-> polychromatic laser (not quite mature yet)

Some challenges of LGS AO

Spot elongation
Sodium layer
is ~10km thick



4m off-axis = 1” elongation
15m off-axis = 4” elongation
-> better to launch from the center of pupil than the edge
-> dynamic refocusing + pulsed laser



Upstream path / diffraction
Laser has to go through turbulence -> LGS is extended
Diffraction from laser launching telescope aperture

-> it is very difficult to create a small size LGS

Spot size excludes some high sensitivity WFS options

Wavefront Sensor Options...

Shack-Hartmann (SHWFS)

Pyramid
- can be fixed (FPYRWFS) or modulated (MPYRWFS)

Curvature (CWFS) (briefly described in this lecture)
- image intra & extra pupil images

Zernike phase constrast (ZWFS)
- focal plane phase mask to encode phase into amplitude

Interferometer
- Pupil plane Mach-Zender interferometer (PPMZWFS)
- Shearing inteferometer

Focal Plane AO (FPWFS) (briefly described in this lecture)

Wavefront Sensor Options...
How to decide ????

Lets start by measuring **how sensitive** these WFSs are...

Sensitivity = how well each photon is used

For a single spatial frequency (OPD sine wave in the pupil plane, speckle in the focal plane):

Error (rad) = Sensitivity / sqrt(# of photons)

IDEAL WFS:

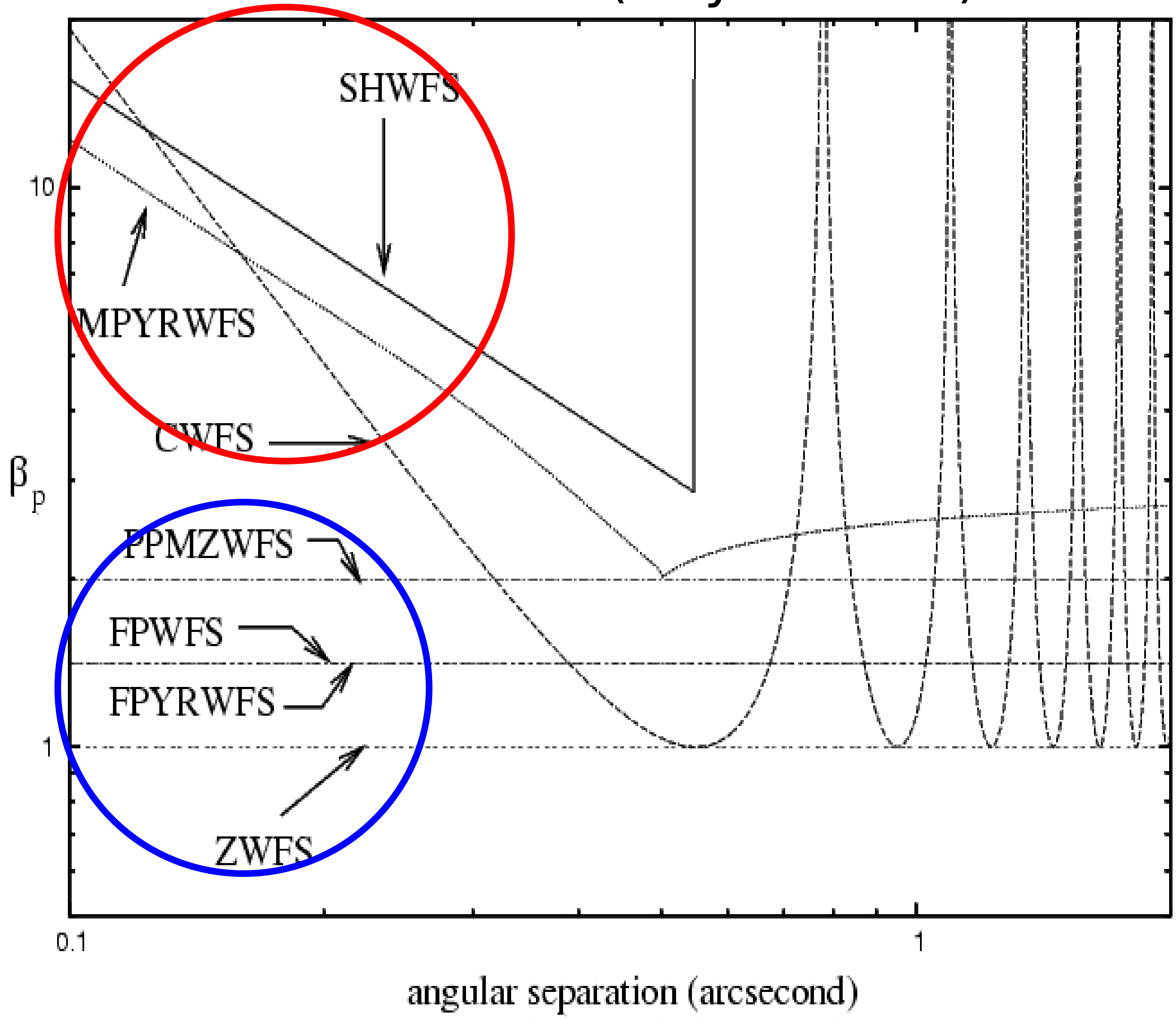
Sensitivity Beta = 1 (1 ph = 1 rad of error)

At all spatial frequencies

Non-ideal WFS:

Beta > 1 (Beta x Beta ph = 1 rad of error)

Wavefront sensors ''sensitivities'' in linear regime with full coherence (Guyon 2005)



Square root of # of photons required to reach fixed sensing accuracy

plotted here for phase aberrations only, 8m telescope. Tuned for maximum sensitivity at 0.5” from central star.

To understand the fundamental difference between BLUE and RED wavefront sensors in previous slides, we will look at one member of each group:

Focal plane WFS

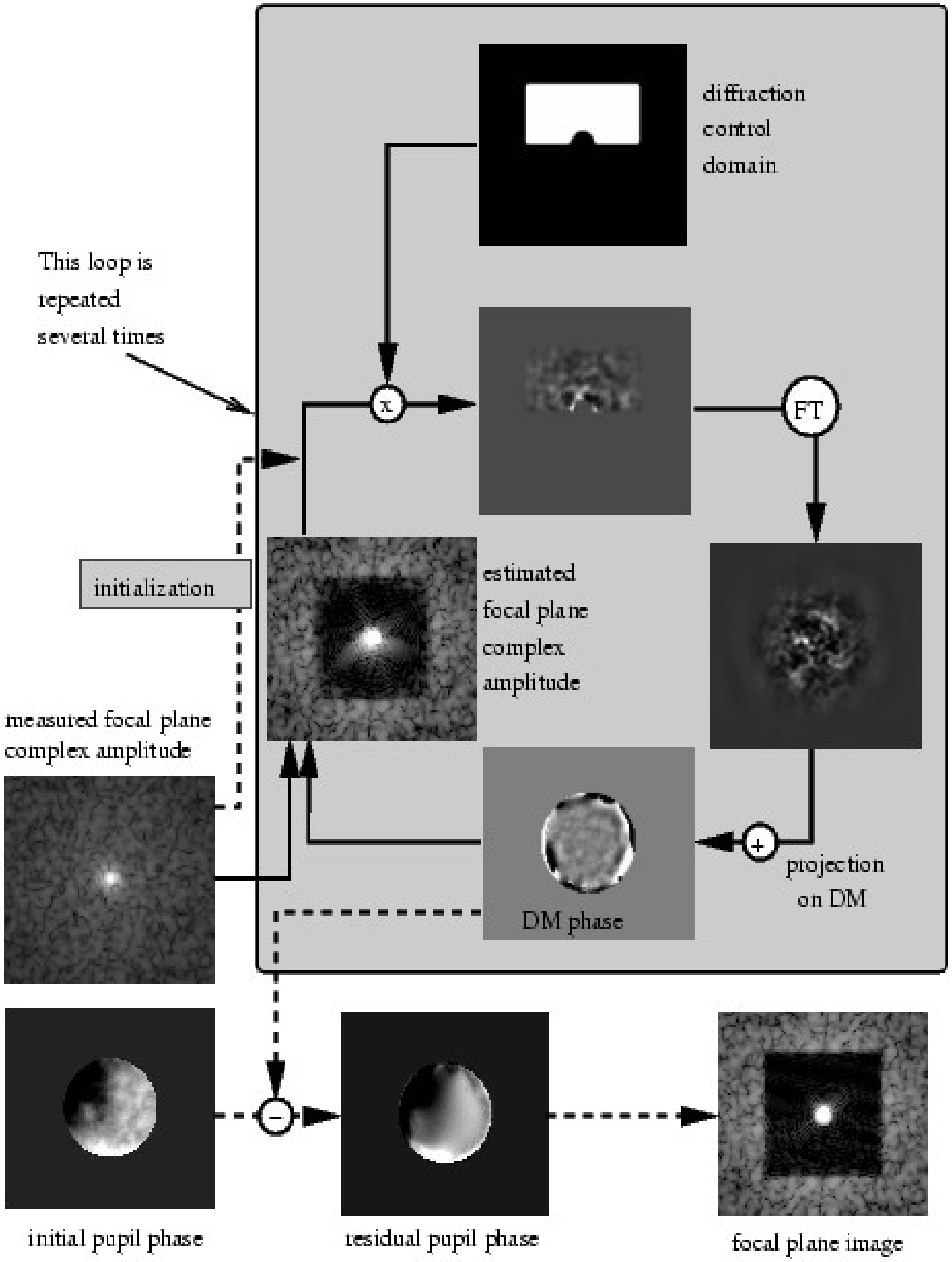
Curvature WFS

Focal plane WFS

If speckle field Complex amplitude is known, DM(s) can be controlled to "perfectly" cancel speckles

DM can be also be asked to create "arbitrary" speckle field for WFS

Malbet, Yu & Shao (1995)
Guyon (2005)
Give'on (2003-2006)
Borde & Traub (2006)



How to **optimally** measure speckle field complex amplitude ?

Use upstream DM to introduce phase diversity.
Conventional phase diversity: focus
With DM: freedom to tune the diversity to the problem

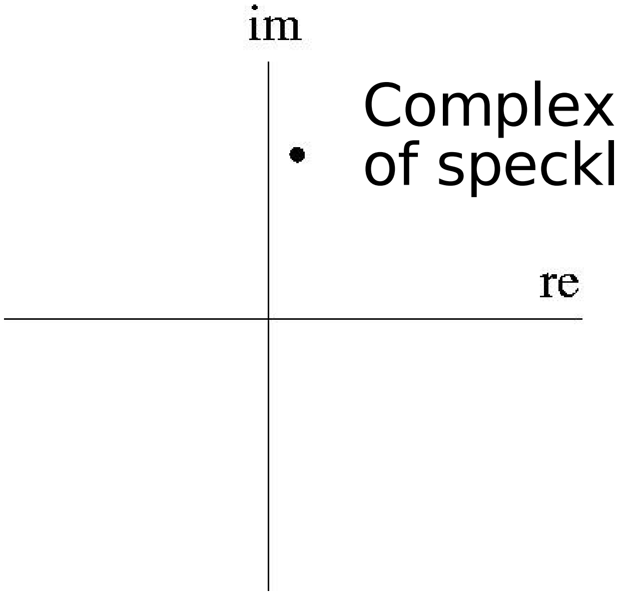
Measure speckle field with no previous knowledge:

- take one frame – this gives a noisy measure of the speckle field amplitude, but not phase

- compute 2 DM shapes which will add known speckles on top of existing speckles. These 2 “additive” speckle field have same amplitude as existing speckles, and the phase offset between the 2 additive speckle fields is PI/2

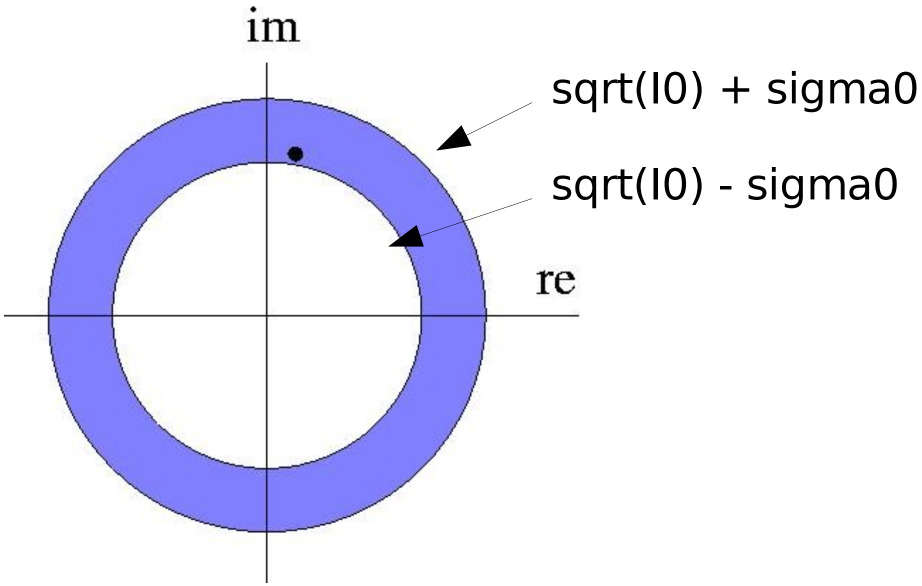
-> for each point in the focal plane, 3 intensities -> single solution for phase & amplitude of speckle field

Initial problem

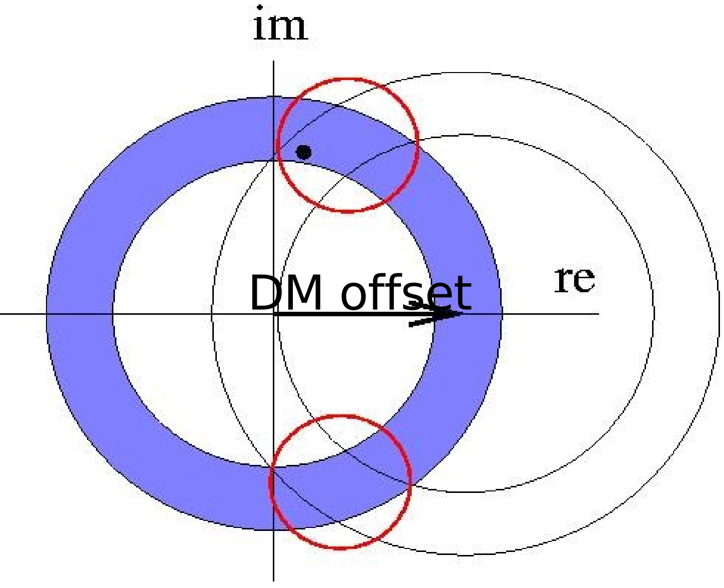


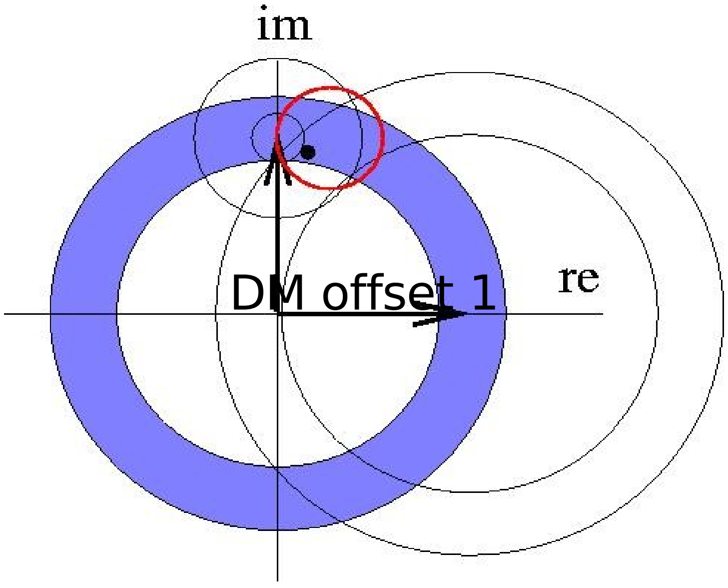
Complex amplitude of speckle

Take a frame -> measured speckle intensity = I0

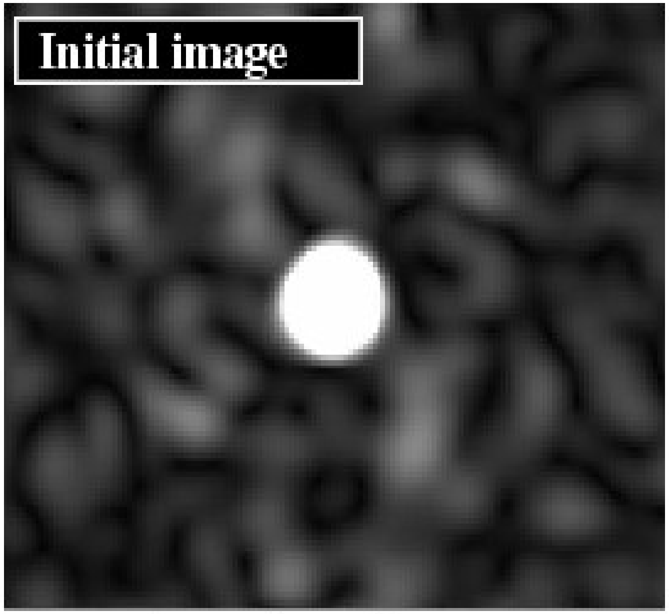


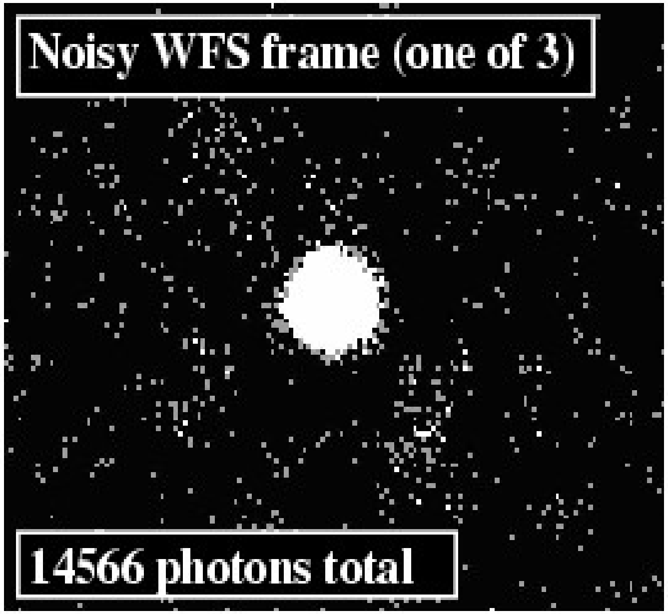
DM offset chosen to be ~ equal to speckle amplitude

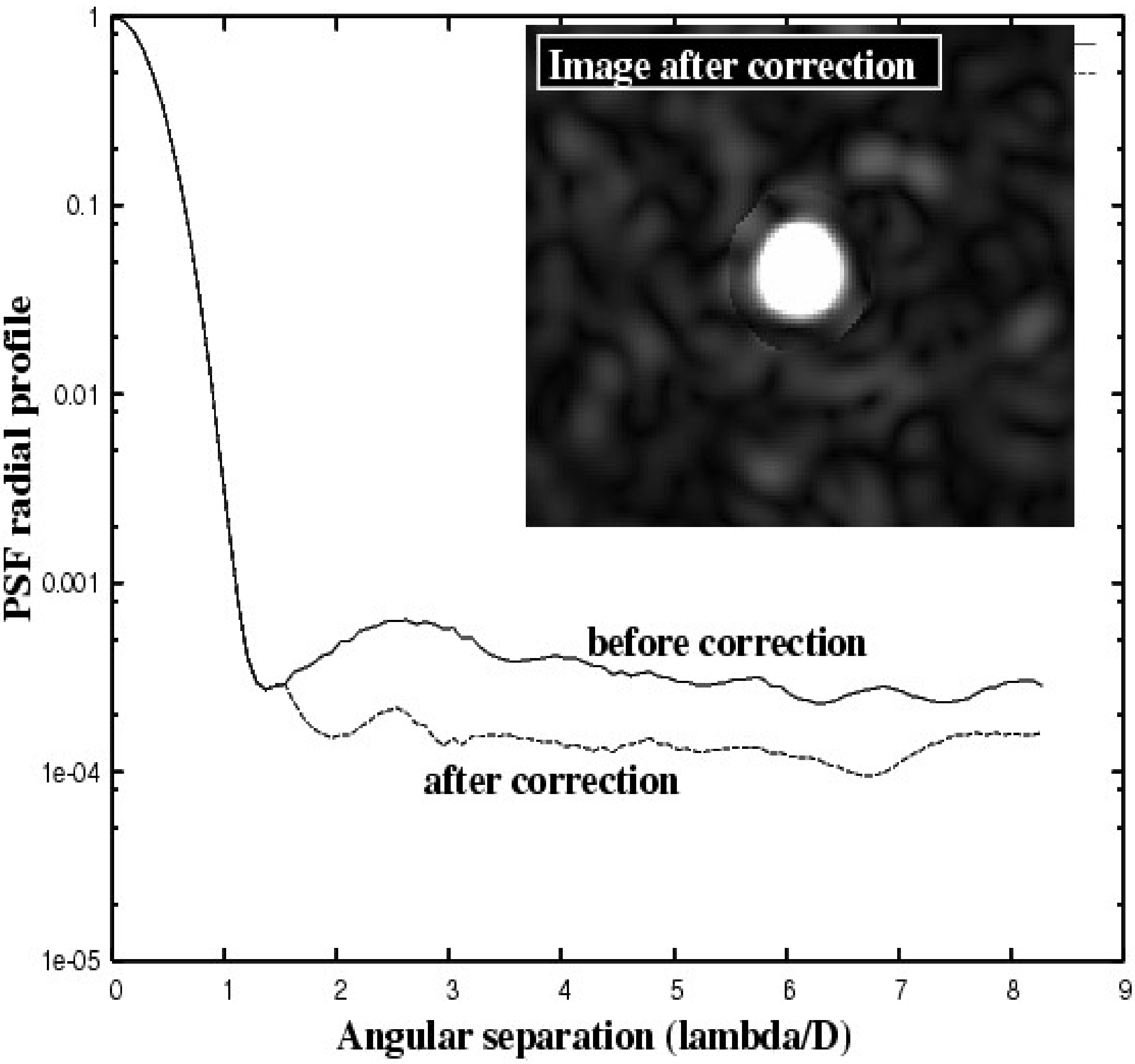




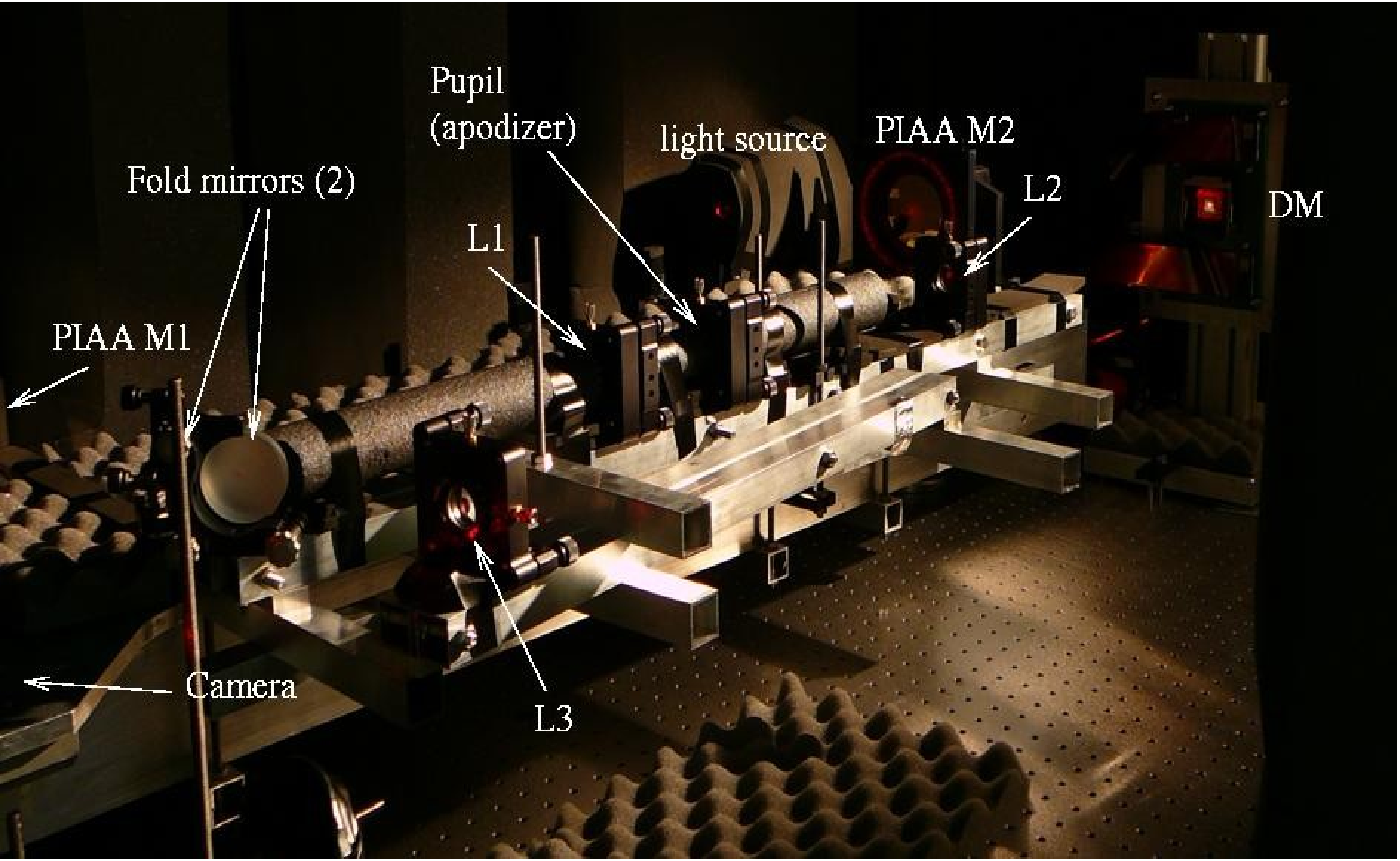
AO correction using Focal Plane WFS: Example simulation



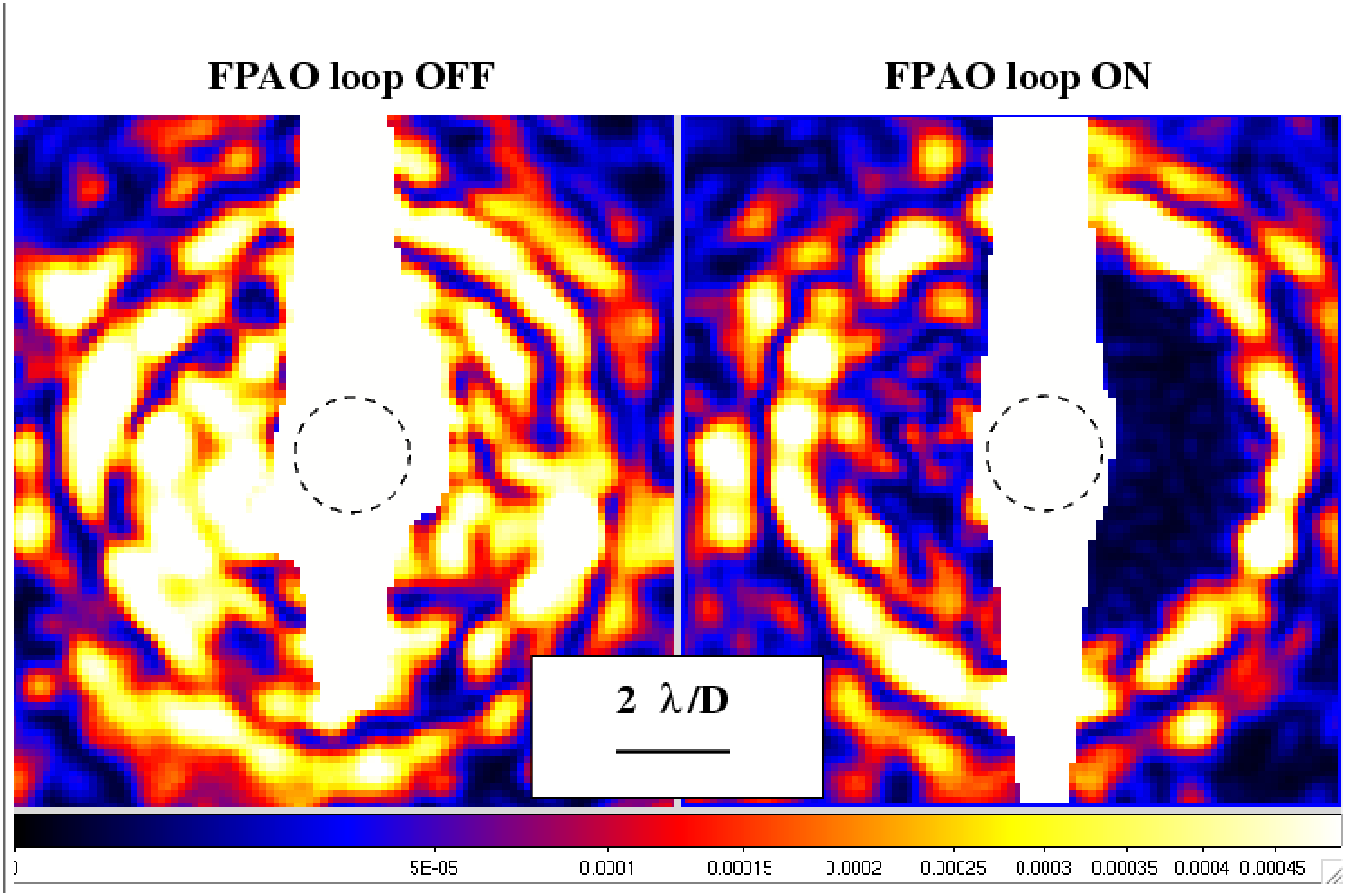




FPWFS is particularly well suited for high contrast Coronagraphic imaging

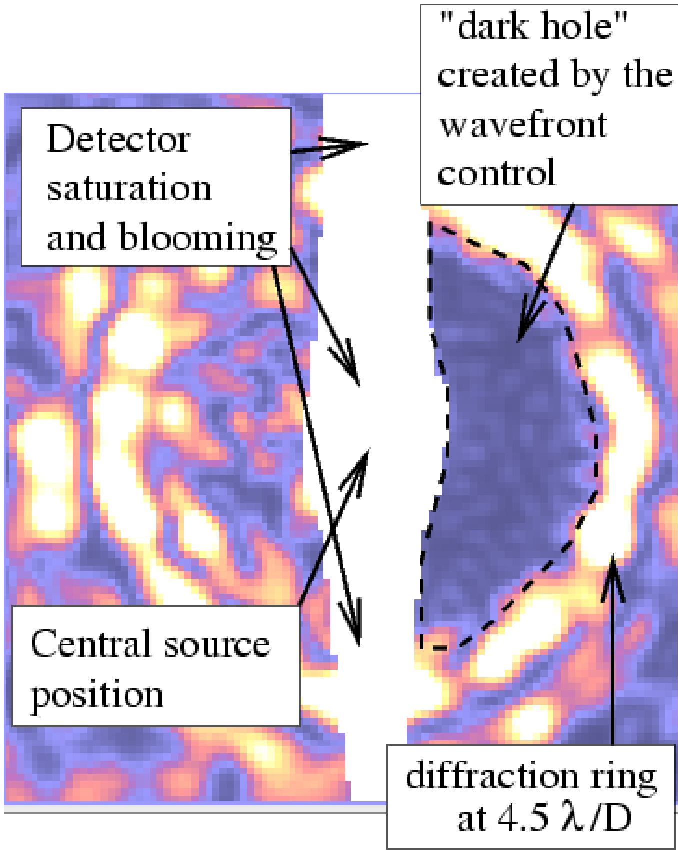


Lab results with PIAA coronagraph + FPAO with 32x32 MEMs DM



FPAO loop OFF

FPAO loop ON



Detector saturation and blooming

"dark hole" created by the wavefront control

Central source position

diffraction ring at 4.5 λ/D

See also results obtained at JPL HCIT & Princeton
So far, these results are obtained at <1 Hz: making FPAO run at ~kHz is challenging (detector, algorithms)

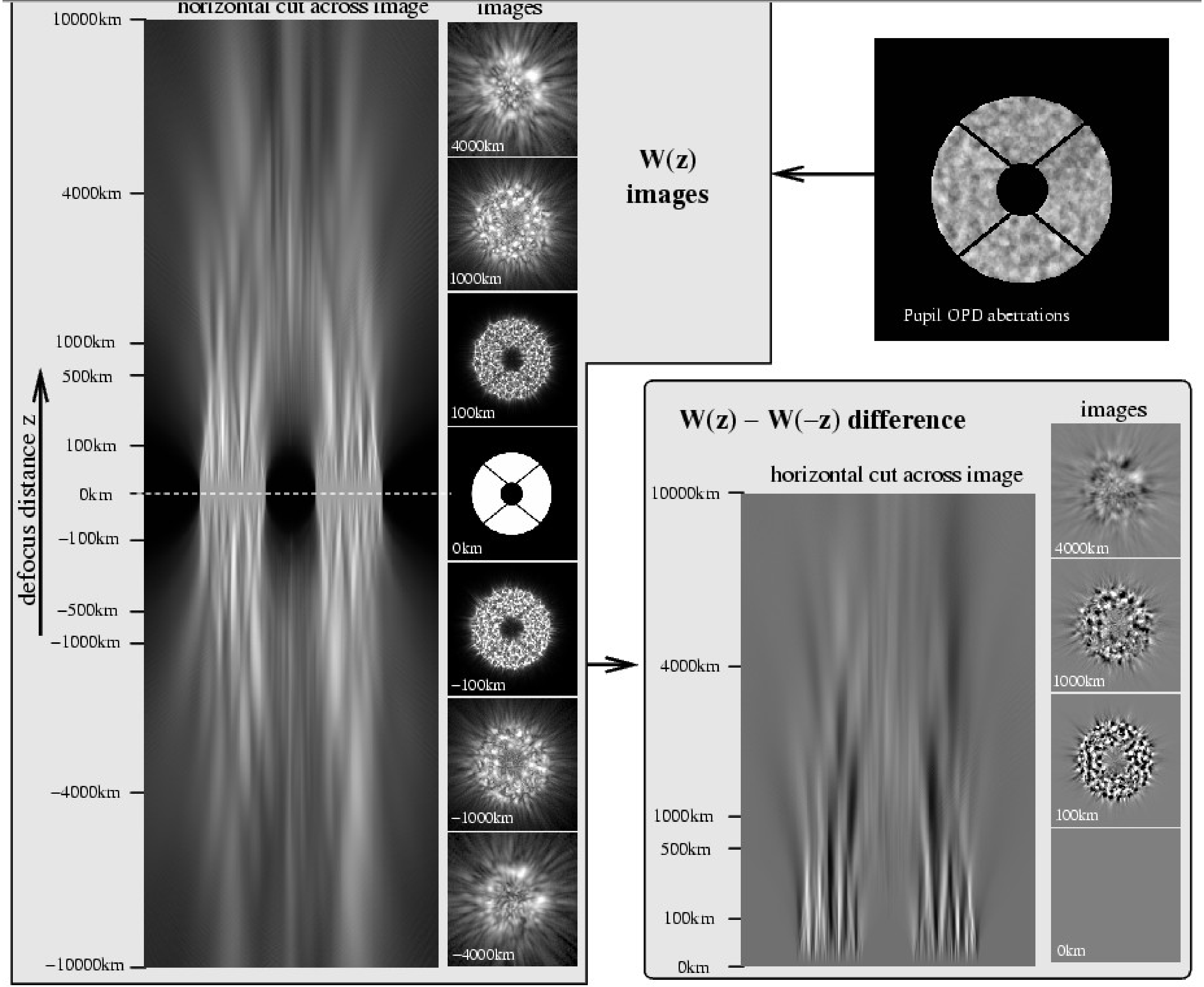
Curvature WFS

Uses **light propagation** to convert phase into intensity
-> measure intensity in at least 2 “defocused” pupil planes and compute phase.

Usually, planes at +dz and –dz, with
dz ~ 1000km are imaged.

If dz “small” (~1000 km), **defocused images are linear function of wavefront curvature**

Next slide shows how phase is converted into intensity modulation in a CWFS

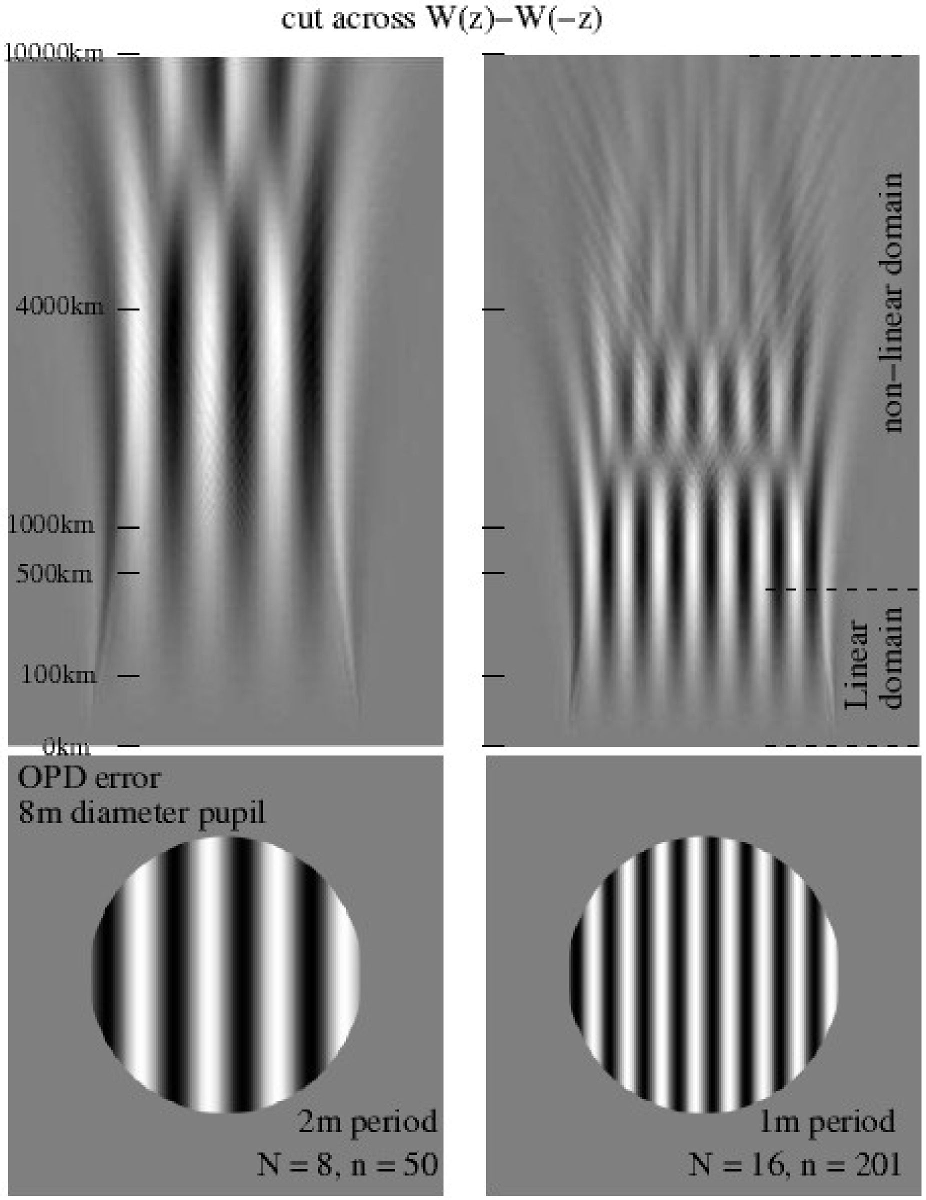

horizontal cut across image
images
10000km
4000km
1000km
500km
100km
0km
−100km
−500km
−1000km
−4000km
−10000km
defocus distance z
4000km
1000km
100km
0km
−100km
−1000km
−4000km
W(z)
images
Pupil OPD aberrations
W(z) – W(−z) difference
horizontal cut across image
images
10000km
4000km
1000km
500km
100km
0km
4000km
1000km
100km
0km

Problem #1:

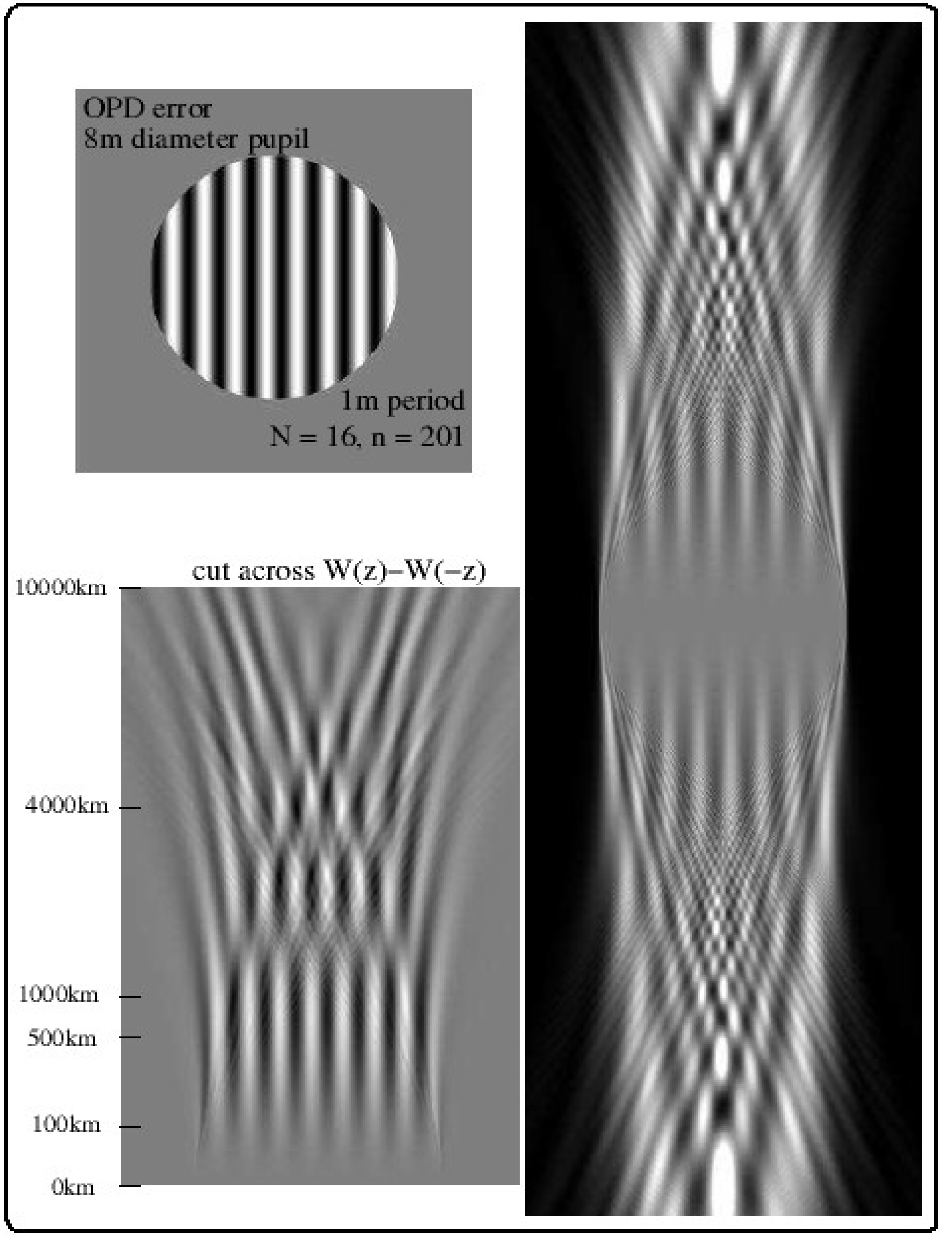
The “Linear” domain of curvature wavefront sensing (= defocus range within which wavefront curvature is linearly transformed into intensity modulation) becomes smaller as the # of actuators increases.

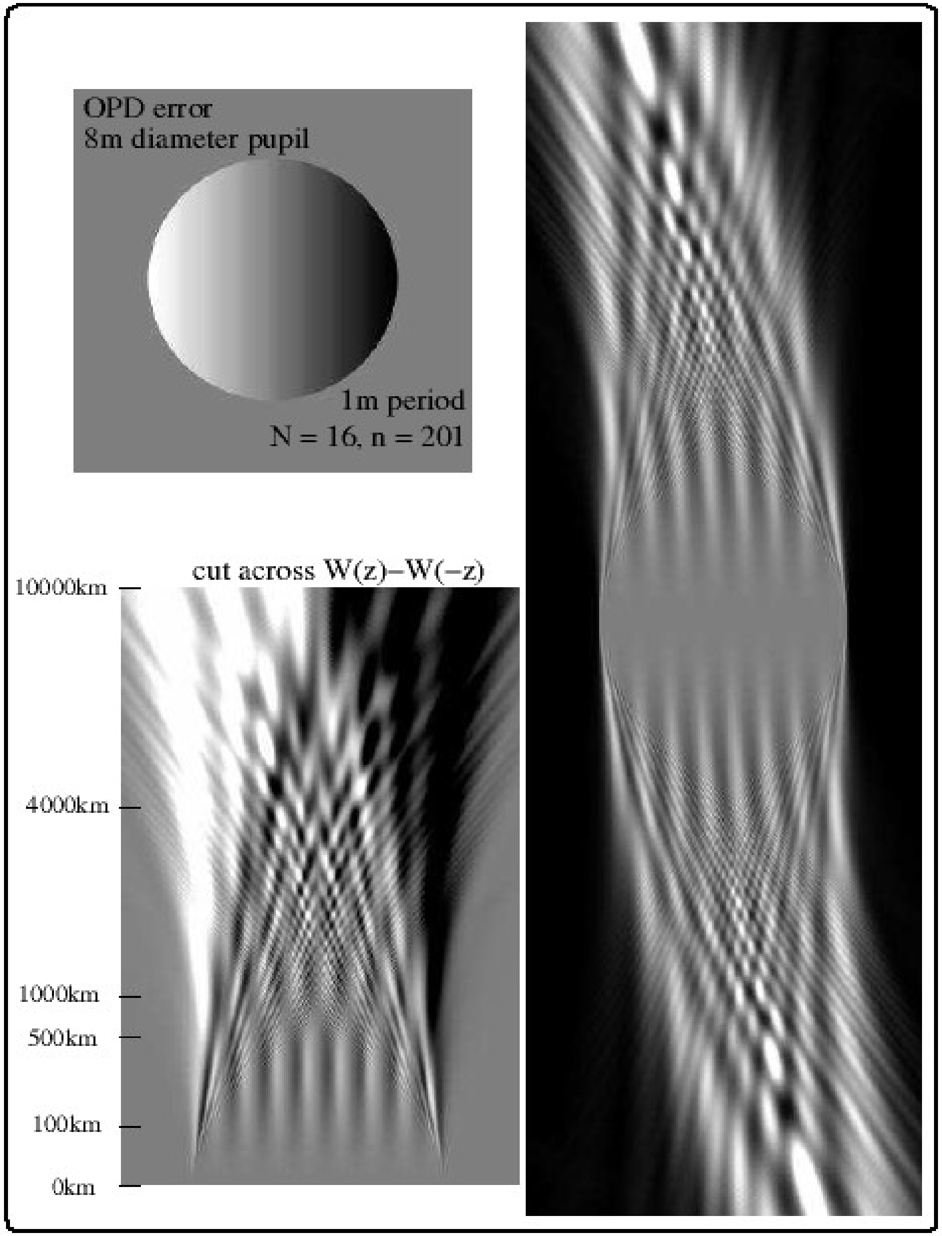
-> defocus distance must be kept small

-> this forces low spatial frequencies to be poorly sensed



Problem #2: Low order aberrations "scramble" high spatial frequencies -> defocus distance must be kept small





Why do SH, Curvature (& modulated pyramid) have bad sensitivity for low order aberrations ?

Good measurement of low order aberrations requires
interferometric combination of distant parts of the pupil
FPWFS does it, but:

- SH chops pupil in little pieces -> no hope !
- Curvature has to keep extrapupil distance small
  (see previous slides) -> same problem

Things get worse as # of actuators go up.
-> **This makes a big difference for ELTs**

Tip-tilt example (also true for other modes):
With low coherence WFS, sigma2 ~ 1/D^2 (more photons)
Ideally, one should be able to achieve:
sigma2 ~ 1/D^4 (more photons + smaller l/D)

This problem is also referred to as <u>"noise propagation"</u> (propagation of photon noise into low spatial frequencies)
Cannot be overcome in SHWFS because pupil is chopped...
However, operation of curvature WFS in non-linear regime, with large defocus distances, solves the noise propagation effect.
Reconstruction algorithm is similar to phase retrieval (algorithm needs to be fast, with few iterations)

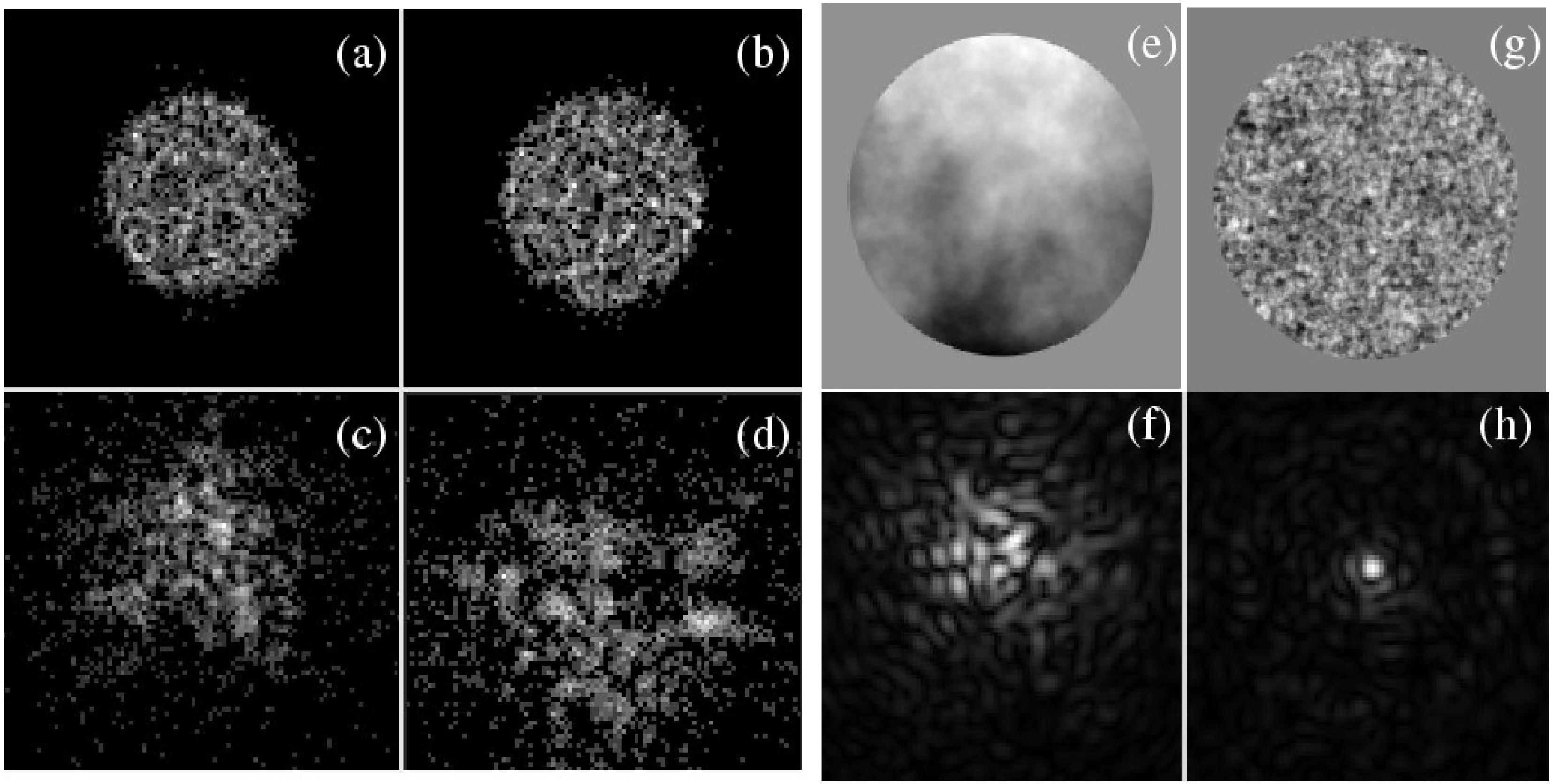


Fig. 9.— Wavefront reconstruction using the algorithm shown in fig. 8. Four noisy defocused pupil images (images (a), (b), (c) and (d)) are acquired to transform the pupil phase aberrations (e) into intensity signals. The input pupil phase is 609 nm RMS, yielding the PSF (f) before correction. After correction, the residual pupil phase aberration (g) is 34.4 nm RMS, allowing high Strehl ratio imaging (h). All images in this figure were obtained at 0.65 μm. The total number of photons available for wavefront sensing in 2e4.

Previous analysis shows that SH, Curv, and Pyramid are very sub-optimal for low spatial frequencies

BUT

To take advantage of high performance WFS options, wavefront needs to be coherent...

Guide “star” for WFS: COHERENCE

COHERENCE = ability to make coherent interferences between different parts of the pupil

Coherence is usually high across small parts of the pupil, low across large parts of the pupil

What makes the guide star “incoherent” ?

Wavefront stability during sampling time

- sampling time too long / turbulence too fast
- sensing wavelength too short
- vibrations

Large time-variable and/or unknown wavefront errors

- poor correction
- open loop wavefront sensing

Angular size of source

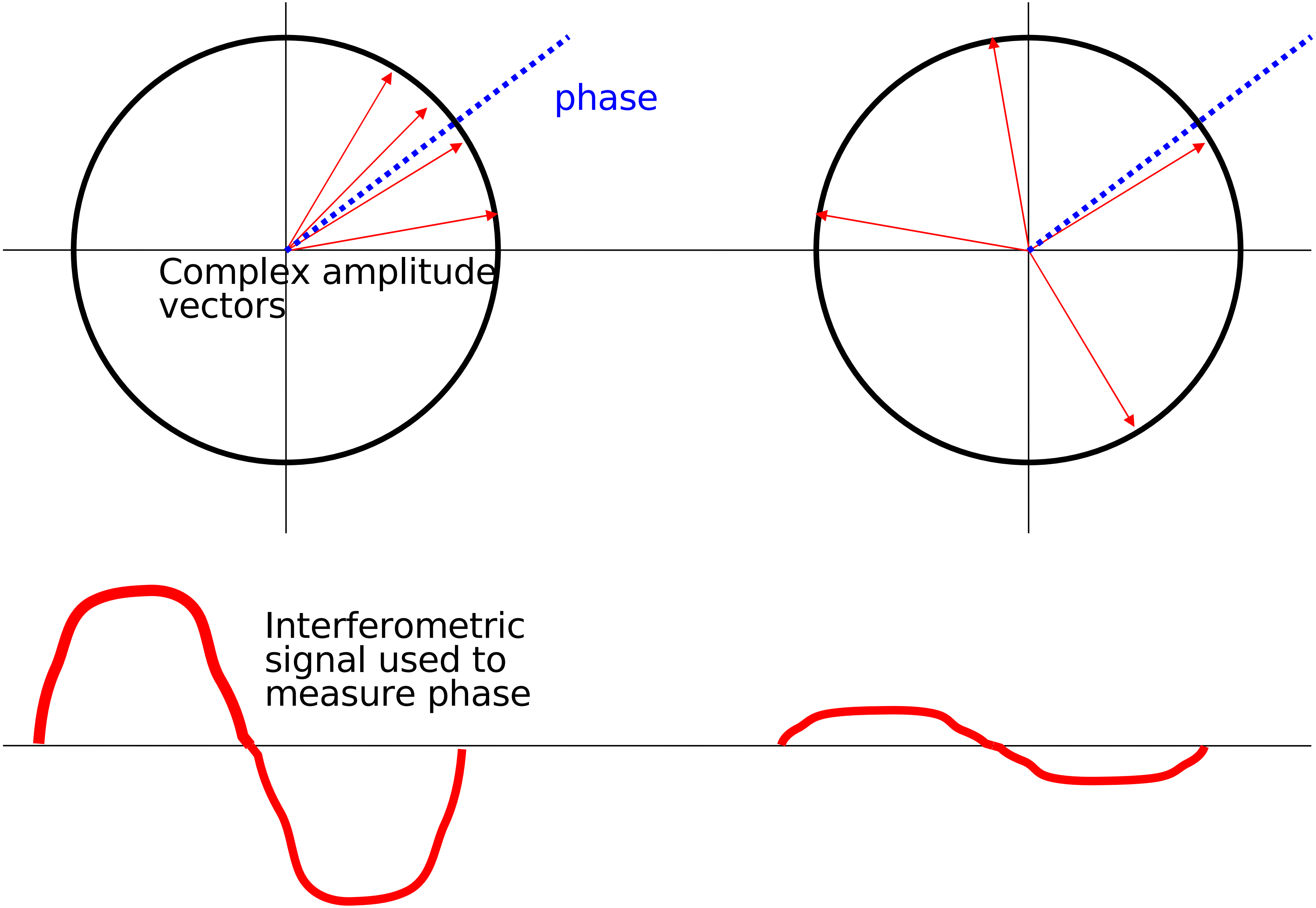
- Atmospheric dispersion
- source resolved > lambda/D

Chromaticity

“interferometer” representation of coherence in WFS

High coherence

Low coherence



phase

Complex amplitude vectors

Interferometric signal used to measure phase

Example of how choosing longer sensing wavelength helps by increasing wavefront coherence (even though phase signal gets smaller !!!)

Closed loop simulations

WFS:
non-linear phase retrieval on curvature wavefront sensor

Same behaviour would be obtained with fixed pyramid

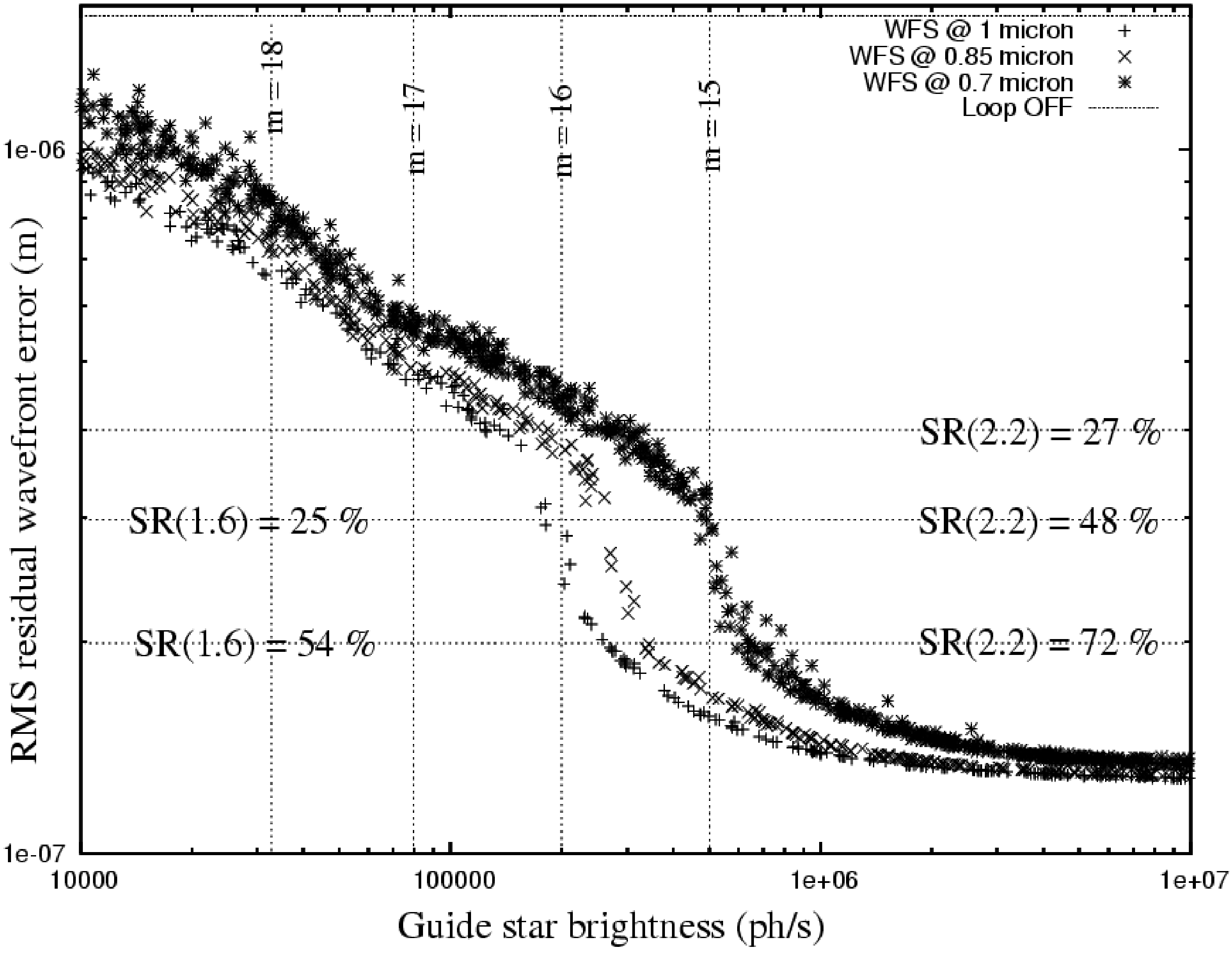
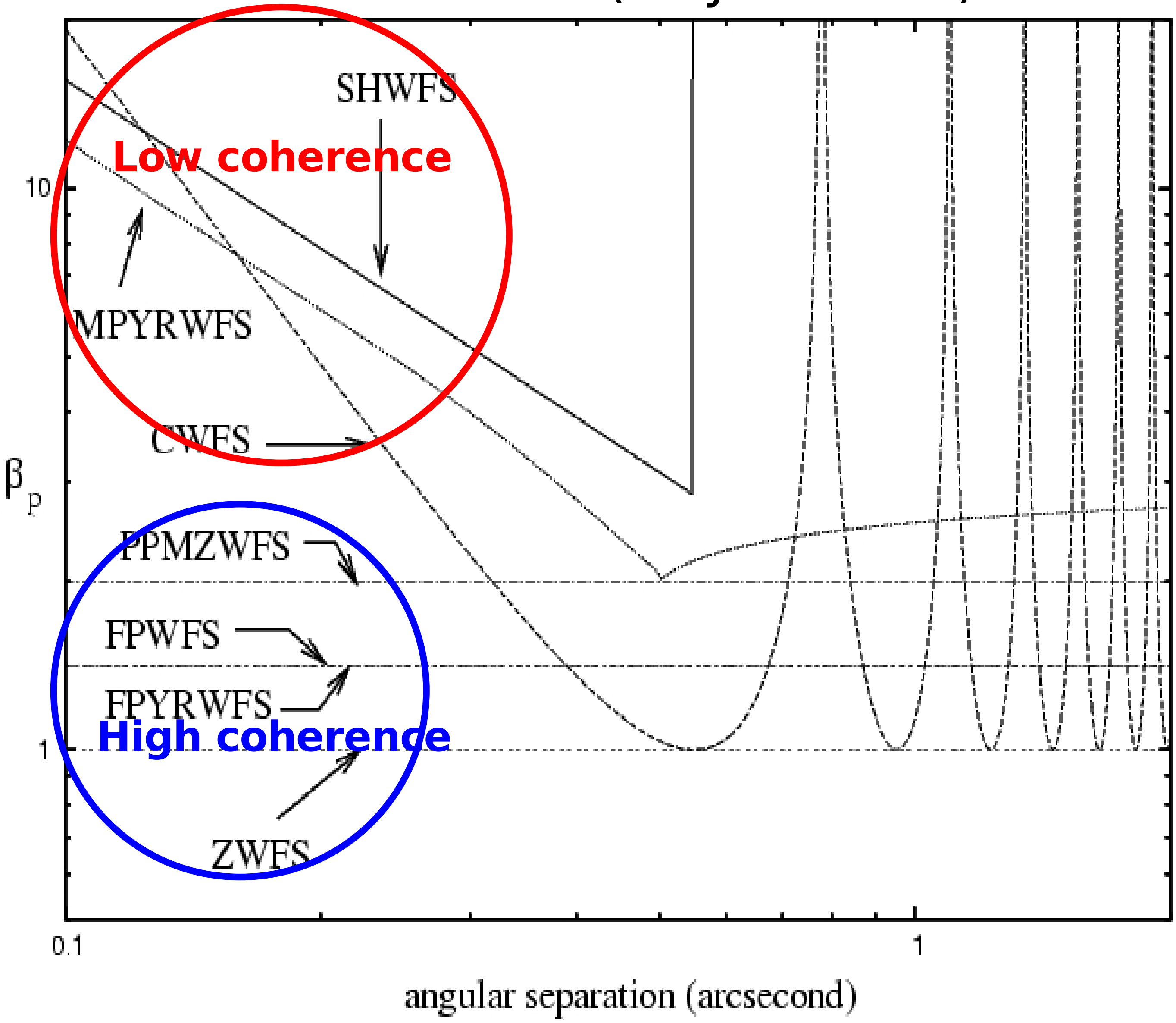


Fig. 11.— Simulated performance of a non-linear Dual stroke Curvature as a function of sensing wavelength $(0.7, 0.85$ and $1.0\ \mu\mathrm{m})$ and guide star brightness. The stellar magnitudes given in this figure assume a 20% efficiency in a $0.5\mu\mathrm{m}$ wide band. See text for details.

Wavefront sensors ''sensitivities'' in linear regime with full coherence (Guyon 2005)



Square root of # of photons required to reach fixed sensing accuracy

plotted here for phase aberrations only, 8m telescope. Tuned for 0.5” separation.

Matching:

Wavefront COHERENCE in WFS **to** **Wavefront sensor**

<< 1 rad

Space Extreme-AO
(Terrestrial Planet Finder)

Second-stage of Extreme-AO
system in near-IR ("Tweeter")

~ 1 rad

Extreme-AO Closed loop in Visible

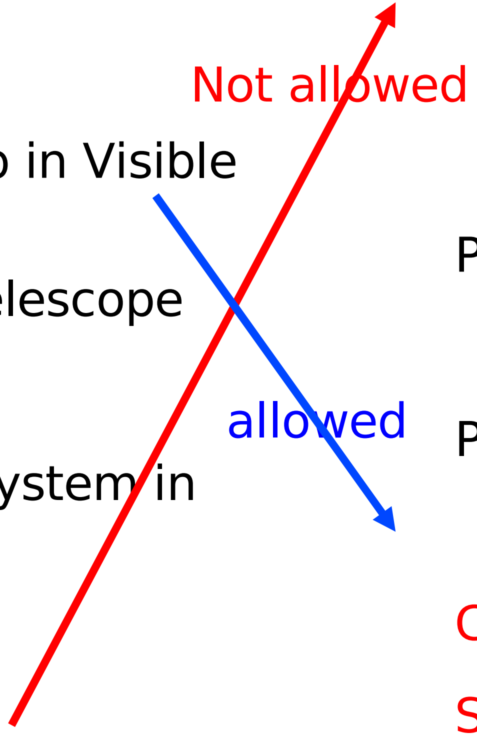
Thermal IR AO on 8m telescope
open loop

"general purpose" AO system in
closed loop

>> 1 rad

LGS AO
GLAO

Open loop AO



Interferometric

Focal plane

Pyramid (fixed)

Pyramid (modulated)

Curvature

Shack-Hartmann

Wavefront sensing options, pros and cons:

SH & conventional Curvature & modulated pyramid

+ robust, large dynamical range, well understood, proven on sky
+ achromatic
SH: + works well on extended sources
Curvature: + few detectors needed
~ non-common path errors can be take out with some work (harder curvature)
- very inefficient at low spatial frequencies (noise propagation)
- large gaps in the pupil can be a problem

Wavefront sensing options, pros and cons:

Pyramid (fixed)

+ no noise propagation

- need <1rad RMS to work well / limited dynamical range

Interferometric

+ no noise propagation

+ works well with gaps

- chromatic, not very robust, problem with extended sources

Focal plane WFS (only good for final stage of high contrast system)

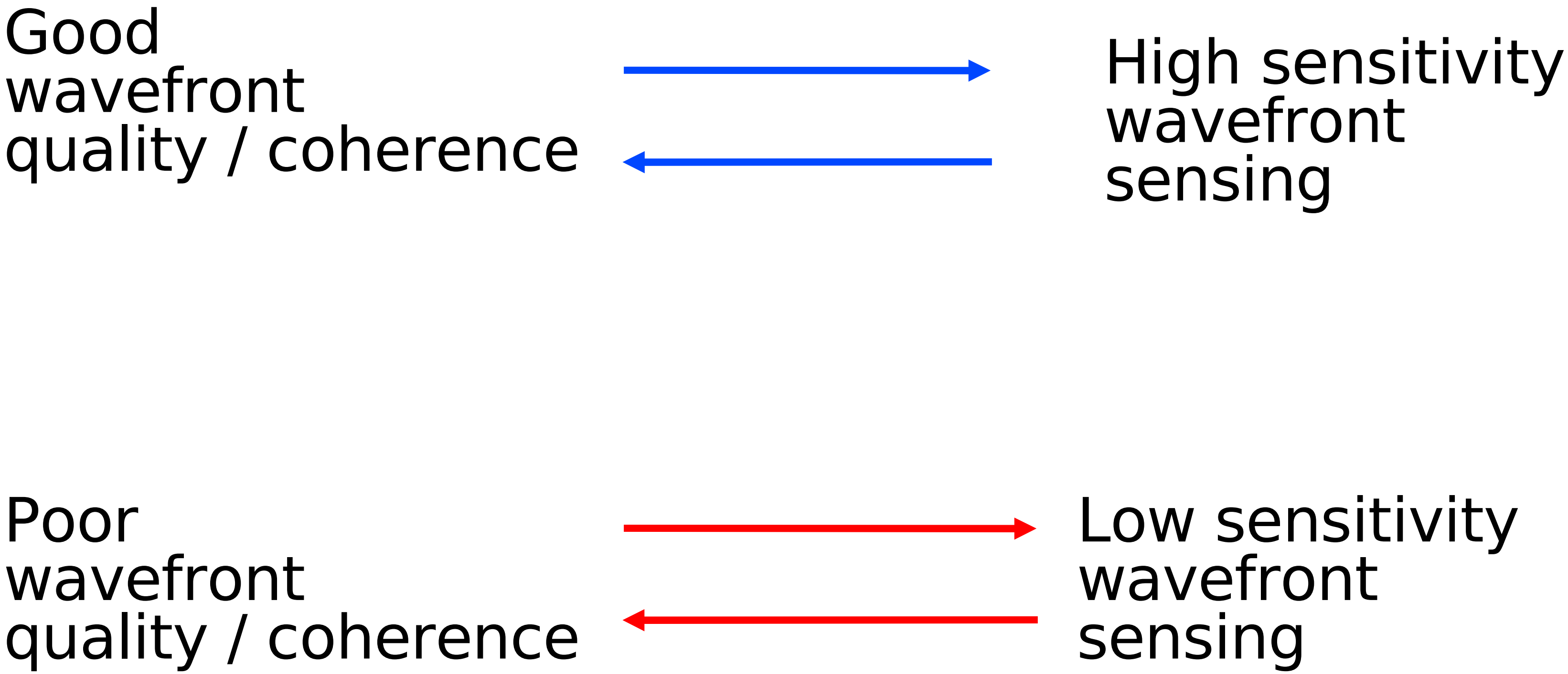
+ no noise propagation, no non-common path errors (even if you use very bad optics) or chromatic errors: **“What you see is what you need to kill”**

- chromaticity vs. FOV

- AO loop will crash if coherence is low

- speed is challenging with near-IR detectors

“The rich get richer, the poor get poorer…”



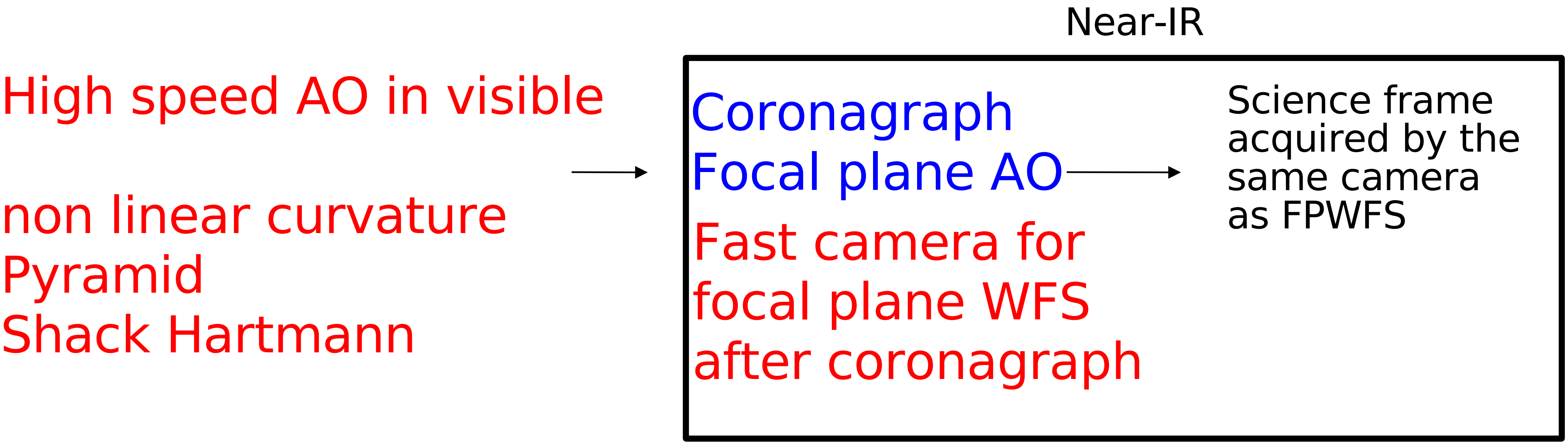
Good
wavefront
quality / coherence

High sensitivity
wavefront
sensing

Poor
wavefront
quality / coherence

Low sensitivity
wavefront
sensing

Example: Possible Coronagraphic ExAO architecture



High speed AO in visible

non linear curvature
Pyramid
Shack Hartmann

Near-IR

Coronagraph
Focal plane AO

Fast camera for
focal plane WFS
after coronagraph

Science frame
acquired by the
same camera
as FPWFS

The first step is used to clean the wavefront within ~ 1 rad in Visible

The second step operates in the high coherence regime, and adopts the FPWFS.

Gemini Planet Imager (GPI) uses a similar strategy, with an interferometer to measure coherent residuals

1. Main challenges / error budget terms in astronomical AO systems

How to design an AO system which meets science requirements & reduces overall error budget ?

- The answer strongly depends on the science objective
  - Solar AO, Extreme AO, MOAO, GLAO, LGS vs. NGS ...

2. Wavefront sensing strategy

- **AO guide star:**
  - **LGS, NGS ? Multiple sources ? Sensing wavelength ?**
- **Choosing the right Wavefront Sensor**

3. From photons to DM commands: making it all work nicely together

AO control

How should the AO system drive the DM from WFS measurements ?

“standard” solution (fast, linear):

- Measure/model how WFS measures DM commands
- If relationship is linear, this is stored as a “response matrix”
“response matrix” is inverted -> “control matrix” (this step usually includes some filtering – see next slide)
- WFS measurements x control matrix = DM commands

This could also be done by computing the wavefront:

WFS measurements -> wavefront -> DM commands

Good AO control allows to separate WFS choice from DM choice:
example: Curvature WFS could run with a MEMs DM

AO control

Modal control/filtering helps a lot

Concept: Run AO loop at different speed for each mode, depending upon mode strength & WFS sensitivity for the mode

- reject "bad modes" which can be produced by DM but not well sensed by WFS
- attenuate known vibrations
- powerful tool for system diagnostic

Example:

mode poorly seen (noisy) by WFS & weak in the atmosphere should be prevented from feeding strong signals to DM

powerful & well sensed mode should be rapidly driving the DM

Modal control continuously tunes the system for optimal perf.

AO system & science instrument

Other ways science instrument can drive AO design:

IR instruments need low thermal background
    -> fewer warm optics
        example: adaptive secondary mirror

Thermal IR instruments may need “chopping” (on source / off source images to calibrate background)
    AO system then needs to be compatible with chopping (this is not easy)

Science instrument can perform its own wavefront sensing
    - This is especially true for Extreme-AO
        Science instrument measure non-common path errors
        example: Focal plane wavefront sensing

Realistic simulation of AO system is extremely useful

AO simulations are relatively accurate, as input and outputs are well known:

- seeing properties are fairly well known (Kolmogorov layers)
- WFS behavior & properties are usually very well known
- Control algorithm identical in simulations & on the sky

AO simulations can investigate:

-> performance vs. # of actuators, DM type/geometry
-> loop instabilities & mode filtering
-> hardware trade-off:
    WFS detector readout noise
    DM hysteresis
    speed of electronics & computer
    Laser power for LGS
    On-axis vs. off-axis LGS
-> alignment tolerance

Telemetry is also very important

Recording WFS and DM data allows:

- seeing estimation & logging
- self-tuning of system
- diagnostics

If a strange behaviour is observed in the AO loop, it is very hard to identify it without being able to "play back" the time when it occurs.

Issues:

Disk space
File management, archiving